Virus Genes (2014) 48:252–259
DOI 10.1007/s11262-013-1030-z

Group B strains of human respiratory syncytial virus in Saudi Arabia: molecular and phylogenetic analysis

Fahad N. Almajhdi · Mohamed A. Farrag · Haitham M. Amer

Received: 3 November 2013 / Accepted: 19 December 2013 / Published online: 27 December 2013
© Springer Science+Business Media New York 2013

Abstract The genetic variability and circulation pattern of human respiratory syncytial virus group B (HRSV-B) strains, identified in Riyadh during the winters of 2008 and 2009, were evaluated by partial sequencing of the attachment (G) protein gene. The second hypervariable region (HVR-2) of G gene was amplified by RT-PCR, sequenced and compared to representatives of different HRSV-B genotypes. Sequence and phylogenetic analysis revealed that all Saudi strains belonged to the genotype BA, which is characterized by 60-nucleotide duplication at HVR-2. Only strains of 2008 were clustered with subgroup BA-IV, while those isolated at 2009 were clustered among the most recent subgroups (particularly BA-X and CB-B). Amino acid sequence analysis demonstrated 18 amino acid substitutions in Saudi HRSV-B strains; among which five are specific for individual strains. Furthermore, two potential N-glycosylation sites at residues 230 and 296 were identified for all Saudi strains, and an additional site at amino acid 273 was found only in Riyadh 28/2008 strain. O-glycosylation was predicted in 42–43 sites, where the majority (no = 38) are highly conserved among Saudi strains. The average ratio between non-synonymous and synonymous mutations (ω) implied stabilizing selection pressure on G protein, with evidences of positive selection on certain Saudi strains. This report provides preliminary data on the circulation pattern and molecular characteristics of HRSV-B strains circulating in Saudi Arabia.

Keywords BA genotype · G protein · Human respiratory syncytial virus group B · Phylogenetic analysis · Saudi Arabia

F. N. Almajhdi (✉) · M. A. Farrag · H. M. Amer
Department of Botany and Microbiology, College of Science, King Saud University, P.O. Box 2455, Riyadh 11451, Saudi Arabia
e-mail: majhdi@ksu.edu.sa

H. M. Amer
Department of Virology, Faculty of Veterinary Medicine, Cairo University, Giza 12211, Egypt

Introduction

Respiratory tract infection in neonates and young children has a substantial health and economic burden on the communities of both developing and developed countries. Human respiratory syncytial virus (HRSV) is the leading respiratory pathogen that causes primary infection in 25–40 % and hospitalization in 0.1–2 % of infants during the first year of life [1]. The virus is also recognized as an important etiology of severe community-acquired respiratory illness in immuno-compromised adults and those suffering from underlying pathology like cardiopulomary disease [2]. As a member of family *Paramyxoviridae*, HRSV is an enveloped, non-segmented, negative single-stranded RNA virus that encodes at least ten viral proteins [3]. Two major groups of HRSV, A and B, were identified using monoclonal antibodies and sequence analyses of several viral protein genes [4, 5].

HRSV has two envelope-associated proteins G and F, which are responsible for cell attachment and entry, respectively. Both proteins are primary targets for host immune response and therefore subjected to a selective evolutionary pressure through accumulation of mutations that alter epitope structure and glycosylation profile [4, 6]. The G protein, in particular, is the most variable viral protein between and within the two groups. This variability is attributed to the eco domain portion, which contains two

hypervariable regions separated by highly conserved cysteine-rich motif of 13 amino acids [6]. Studies of virus diversity have elucidated that strain-specific epitopes are mostly located in the second hypervariable region (HVR-2), and that this domain reflects the variability of the entire genome [5, 7]. Hence, the majority of phylogenetic studies worldwide utilized HVR-2 for genotype classification of HRSV strains [8–10].

Genetic diversity studies among HRSV-B strains has received less concern as compared to HRSV-A strains due to: (i) predominance of HRSV-A in most epidemics, (ii) higher variability and rapid evolution of group A viruses [11], and (iii) infection with HRSV-A is profound causing more severe disease form [12]. Genetic studies have demonstrated the existence of nine distinct genotypes of HRSV-B: GB 1–4 [8, 13], SAB1-4 [14, 15] and BA [16]. The later genotype was identified at 1999 and is characterized by 20-amino acid duplication at the HVR-2 of G protein. Evidences have been accumulated that BA genotype is going to completely replace other HRSV-B genotypes and to be the sole genotype circulating globally. However, sequence divergence among BA genotype strains has been reported as a result of accumulation of point mutations in the duplicated region. Consequently, BA genotype has been subdivided into 10 subgenotypes (BA I-X) [17], besides the novel subgenotype CB-B detected in Korea [10].

Evidences for co-circulation of both HRSV groups in Saudi Arabia were demonstrated only recently [18, 19]. During the winter seasons of 2007/08 and 2008/09, HRSV was identified in nasopharyngeal aspirates collected from hospitalized children in Riyadh. Viruses of both groups were propagated in cell culture and were characterized antigenically. Since group A viruses were the most predominant among positive samples, they received the primary concern regarding analysis of their molecular characteristics and antigenic diversity [19]. In this report, sequence analysis was extended to study the variability of HRSV-B strains circulating in Riyadh during both winter seasons and their phylogenetic relationship with the circulating strains worldwide.

Materials and methods

Virus strains

The HRSV strains examined in this study were recovered from hospitalized children with acute lower respiratory tract infection at King Khalid University Hospital, Riyadh, Saudi Arabia during the winters of 2007/08 and 2008/09. Clinical samples were obtained, after getting an informed written consent from the guardians of the patients, using protocols approved by the Ethical Committee of King Saud University. Virus identification and typing was performed using RT-PCR. HRSV was isolated from clinical samples in HEp-2 cell culture and the virus identity was confirmed using ELISA, immunofluorescence and neutralization assays [19]. A total of 16 HRSV-B strains (representing 41 % of all HRSV positive samples) were characterized and utilized in the current analyses. The nomenclature used for describing these Saudi strains indicates the place of isolation, laboratory code number and the epidemic season (e.g. Riyadh 28/2008).

RT-PCR and sequencing

RNA was isolated from the virus strains using QIAamp viral RNA extraction kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol. The oligonucleotide primers used for RT-PCR were designed, to amplify HVR2 of G protein gene, according to the published sequence of the prototype strain 18537 [6]. These primers are: HRSV-G-F2 (5′-CAAGATGCAACAAGC CAGATC-3′) and HRSV-G-R2 (5′-ACTGCACTGCATGT TGATTG-3′). RT-PCR was conducted in a total volume of 50 µl containing 1 µg RNA, 0.6 µM each primer and 1× FideliTaq RT-PCR master mix plus (GE Healthcare, Buckinghamshire, UK). The reaction tubes were incubated, in GeneAmp 9700 thermal cycler (Applied Biosystems, Foster City, CA), at 42 °C for 30 min (reverse transcription), 94 °C for 3 min (initial denaturation), followed by 40 cycles of denaturation at 94 °C for 30 s, annealing at 55 °C for 30 s and extension at 68 °C for 1 min, and lastly 68 °C for 10 min (final extension). The amplified products were separated by electrophoresis in 1.5 % agarose gel and were purified with Illustra™ GFX PCR DNA and Gel Band Purification Kit (GE Healthcare). Sequencing of the purified products on both strands was conducted using the specific PCR primers at GTAC Biotech (Cologne, Germany). Sequence contigs were edited and assembled using Bioedit program, version 7.0 (Ibis Biosciences, Carlsbad, CA).

Analysis of nucleotide and deduced amino acid sequence data

Sequences of Saudi HSRV-B strains (both nucleotides and deduced amino acids) were aligned using Clustal W, available as an embedded function in MegAlign program of Lasergene software, version 3.18 (DNAStar, Madison, WI). Sequence polymorphisms were recorded by comparison with sequences collected from GenBank that represent strains of different HRSV-B genotypes (Table 1). The number and distribution of potential N- and O-glycosylation sites were predicted within the deduced amino acid

Table 1 List of HRSV-B strains included in sequence and phylogenetic analysis

Strain	Country	Collection date	Accession no.	Reference
Riyadh 28-2008	Saudi Arabia	2008	JF714707	This study
Riyadh 52-2008	Saudi Arabia	2008	KC791694	This study
Riyadh 85-2009	Saudi Arabia	2009	KC791695	This study
Riyadh 86-2009	Saudi Arabia	2009	JF714708	This study
Riyadh 87-2009	Saudi Arabia	2009	KC791696	This study
Riyadh 133-2009	Saudi Arabia	2009	KC791697	This study
BA-100-04	Argentina	2004	DQ227365	[28]
BA-1161-02	Argentina	2002	DQ227377	[28]
BA-1271-02	Argentina	2002	DQ227380	[28]
BA-1607-04	Argentina	2004	DQ227397	[28]
BA-3931-99	Argentina	1999	DQ227365	[28]
BA-4862-03	Argentina	2003	DQ227389	[28]
BA-4974-03	Argentina	2003	DQ227392	[28]
BA-5140-03	Argentina	2002	DQ227393	[28]
BE-13159-02	Belgium	2002	AY751117	[24]
Cam2009-2145	Cambodia	2009	JN119987	[14]
Cam2009-2164	Cambodia	2009	JN119989	[14]
HR-2085-06	Croatia	2006	AB603478	[31]
NG-004-03	Japan	2003	AB175819	[32]
NG-006-03	Japan	2003	AB175820	[32]
NG-013-07	Japan	2007	HM459873	[33]
NG-065-07	Japan	2007	HM459891	[33]
Ken-109-02	Kenya	2002	AY524573	[26]
CB89-08	Korea	2008	AHQ699308	[10]
CB181-08	Korea	2008	HQ699295	[10]
CB757-09	Korea	2009	HQ699304	[10]
CB1119-10	Korea	2010	HQ699289	[10]
SA0025	South Africa	2001	AF348825	[15]
SA0028	South Africa	2001	AF348812	[15]
SA97D934	South Africa	1997	AF348817	[15]
SA98D1656	South Africa	1998	AF348826	[15]
SA98V192	South Africa	1998	AF348811	[15]
SA99V800	South Africa	1999	AF348821	[15]
MOZ-202-99	Mozambique	1999	AF309677	[35]
MON-4-01	Uruguay	2001	AY488801	[36]
18537	USA	1962	M17213	[6]
AL19734-4	USA	2000	AF233924	[8]
B1	USA	1997	AF013254	[37]
CH10b	USA	1998	AF065250	[13]
CH93-9B	USA	1998	AF065251	[13]
CN1839	USA	2000	AF233926	[8]
NY01	USA	2000	AF233931	[8]
NY97	USA	2000	AF233932	[8]
WV15291	USA	1991	M73542	[38]

sequences using NetNglyc 1.0 (http://www.cbs.dtu.dk/services/NetNGlyc) and NetOglyc 3.1 (http://www.cbs.dtu.dk/services/NetOGlyc), respectively. Synonymous and non-synonymous mutations were evaluated by the method of Nei and Gojobori [20] using SNAP program (www.hiv.lanl.gov) [21].

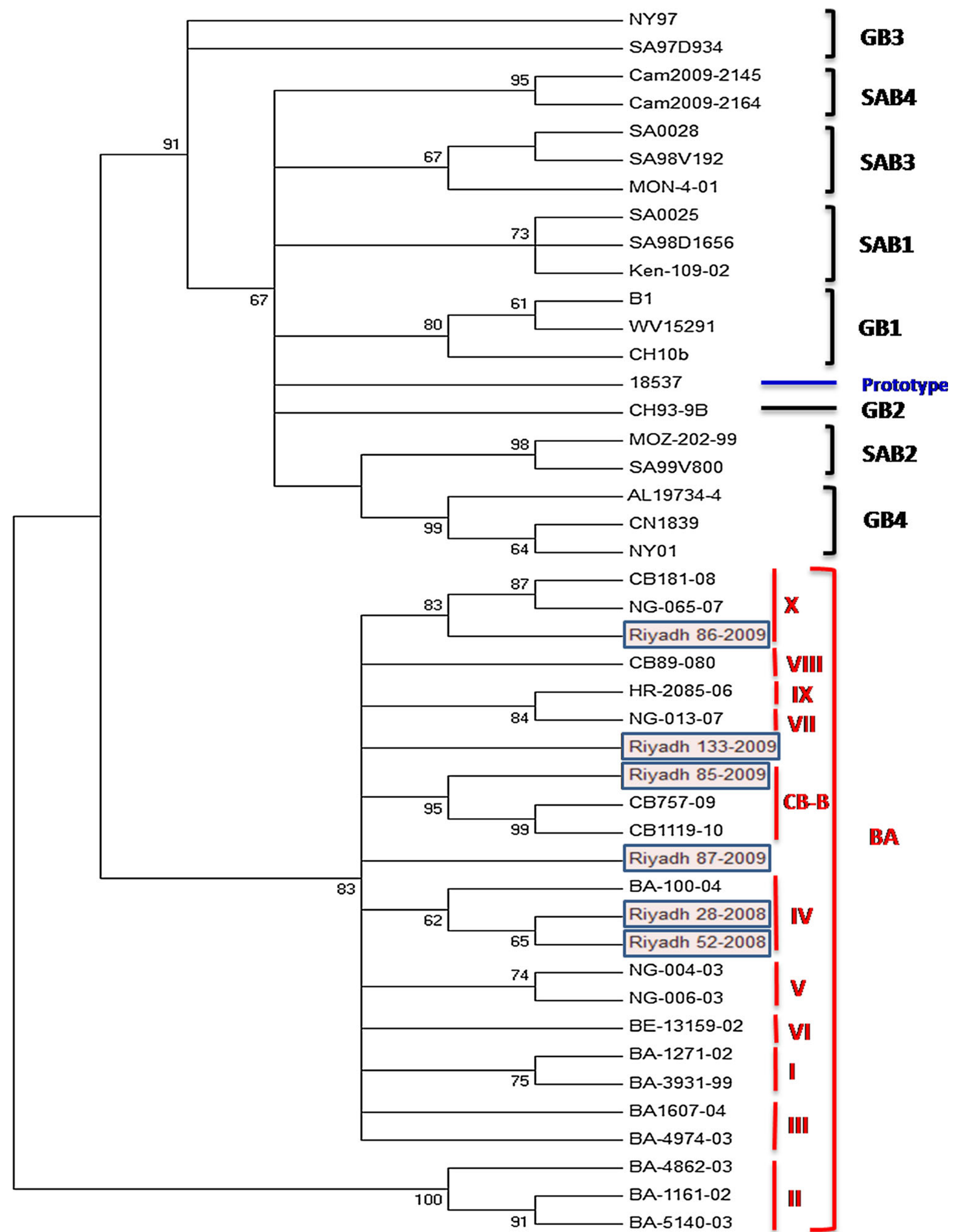

Fig. 1 Phylogenetic tree for HRSV-B nucleotide sequences from HVR-2 of G protein gene. Sequences were aligned using Clustal W and the phylogram was generated by the neighbor-joining method using MEGA5.1 program. The numbers at the internal nodes of the tree represent the *bootstrap values* of 1,000 replicates. Only values exceeding 60 % are shown. Genotypes are shown at the *right side* in *brackets*. The prototype strain 18,537 was used as an outgroup and marked in *blue*. BA genotype and the belonging subgroups are denoted by *red*. Saudi strains are indicated by *shaded squares* (Color figure online)

Construction of phylogenetic tree and genotype distribution

The nucleotide sequence of C-terminal region of HVR2 of G protein gene was utilized for phylogenetic analysis. This region spanned bases 637–879 (243 nt) of prototype strain 18,537 and 637–948 (312 nt) of strain BA-3931-99 that represents BA genotype with 60 nt duplication. The phylogenetic tree was constructed using the neighbor-joining method with Molecular Evolutionary Genetics Analyses (MEGA) software, version 5.1 [22]. Bootstrap probabilities of 1,000 replicates were calculated to evaluate the tree robustness. Only values that surpassed 60 % were indicated at the branch notes. Unique Saudi HRSV-B sequences were only included in tree construction. For comparative and genotyping purposes, the phylogenetic analysis included sequences of 38 strains isolated from different countries worldwide and available in GenBank (Table 1). These strains were assigned previously to the different HRSV-B genotypes and subgenotypes according to the literature.

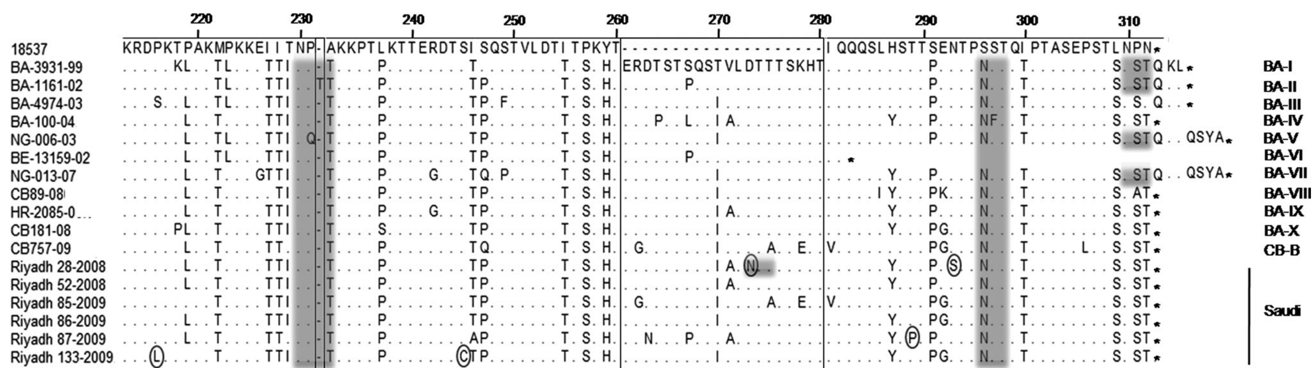

Fig. 2 Deduced amino acid sequence alignment of G protein HVR-2 of HRSV-B Saudi strains. *Alignment* is shown relative to the prototype strain 18,537. Representatives of the different subgroups of BA genotype are included. Identical residues are indicated by *dots* and sequence variation is denoted by a *single letter* code, while gaps are *dashed*. Stop codons are indicated by *asterisks* and Saudi-specific residues are *encircled*. Potential N-glycosylation sites (NXT; X is any amino acid except prolin) are *shaded*. The amino acid insertion and duplication are shown in *boxes*

Nucleotide sequence accession numbers

The nucleotide sequences of HRSV-B strains analyzed in this study are available with the following accession numbers [GenBank: JF714707, JF714708 and KC791694-KC791697].

Results

Phylogenetic analysis and genotype distribution

Among 16 Saudi HRSV-B strains, only six with unique sequences were utilized for the analysis of the phylogenetic relationship and circulation pattern. Since most of the published sequences included the C-terminal region of HVR-2 of G protein gene (last 213–324 bases; according to the protein length), this fragment was only considered in this analysis study. All of the Saudi HRSV-B strains belonged to the newly emerged and rapidly expanding genotype BA that is characterized by 60 nucleotide duplication at HVR-2 of G protein gene. Two of these strains were clustered with subgroup BA-IV, one with BA-X and one with CB-B; the bootstrap values ranged from 62 to 95 %. Two Saudi strains (Riyadh 87/2009 and Riyadh 133/2009) were not assigned to a distinct BA subgroup (Fig. 1). Temporal clustering of Saudi HRSV-B strains was also recognized. For instance, strains recovered during 2008 were only clustered with subgroup BA-IV, while those of 2009 were mostly correlated with the recently identified subgroups (BA-VII to BA-X, and CB-B). Sequence analysis of the analyzed fragment further potentiated this relationship where the homology between Saudi strains and BA genotype was the highest (94.4–96.4 and 89.1–93.1 % at the nucleotide and deduced amino acid levels, respectively).

Molecular characterization of C-terminal variable region of G protein

Protein length and stop codon usage

The partial G protein sequence of Saudi strains was compared to representatives of all HRSV-B genotypes for determination of their distinct molecular features. Although a significant variation was observed in the protein length among and within HRSV-B genotypes (282–319 amino acids), all Saudi strains possessed a single protein length of 312 amino acids (Fig. 2). Similarly, among three alternative types of stop codon observed in HRSV-B strains, Saudi strains utilized only TAA codon. These common characteristics of protein length and stop codon usage were also identified in other members of BA subgroups relevant to Saudi strains (e.g. IV, X, and CB-B) (Fig. 2).

Amino acid sequence diversity

Analysis of the amino acid sequence variation among different HRSV-B genotypes revealed that amino acid substitution, insertion and duplication are the commonly observed changes. Several genotype-specific amino acid substitutions were identified (e.g. P223T for GB1, N230D and Q248R for GB2, S278F for SAB1, K223E and T256A for SAB2, and S288F for SAB4). A single amino acid insertion between residues 231 and 232 was only recognizable in BA-II subgroup. The BA genotype was characterized by 20-amino acid duplication between residues 241 and 281. The duplicated segments were not completely

identical and there is no correspondence between the accumulated mutations in both fragments (Fig. 2).

As mentioned before, all Saudi strains were classified as members of the genotype BA with a typical 20-amino acid duplication. Eighteen amino acid substitutions were recorded in Saudi strains, where Riyadh 52/2008 and Riyadh 86/2009 strains did not show any changes, and Riyadh 85/2009 and Riyadh 87/2008 strains displayed the highest number (six) of amino acid substitutions. Certain amino acids were recognized to be specific for individual Saudi strains such as: D273 N and N293S for Riyadh 28/2008, T289P for Riyadh 87/2009, and P216L and S245C for Riyadh 133/2009 (Fig. 2).

Glycosylation sites

Two N-glycosylation sites were predicted at the C-terminal region of G protein of Saudi strains using Net-O-glyc 1.0 program (Fig. 2). The first, located at residue 230, was conserved among all HRSV-B genotypes except GB4, while the other, at residue 296, was less conserved since it is absent in few members of genotypes GB3, SAB4, BA and the prototype strain 18537. An additional N-glycosylation site at amino acid 273 was observed solely in Riyadh 28/2008 strain as a result of the unique amino acid substitution aspartic acid (D) to asparagine (N). It is worth to mention that the later site is situated in the second segment of the 20- amino acid duplication of this Saudi strain.

In contrary, the number of O-glycoylation sites in C-terminal region was enormous due to the abundance of serine and therionine residues [23]. Using Oligo-O-glyc 3.1 program, 42–43 potential O-glycoylation sites were predicted in Saudi strains with G-scores exceeding 0.5. The majority of these sites (no = 38) were absolutely conserved; among which 9 and 8 were located in the two copies of 20-amino acid duplication, respectively. Riyadh 87/2009 strain was characterized by the lack of three potential O-glycosylation sites (residues 246, 267 and 289), and the existence of two extra-sites at residues 229 and 270. Riyadh 85/2009 and Riyadh 133/2009 strains had additional sites at residues 275 and 245, respectively.

Synonymous and non-synonymous mutation rate

To evaluate the evolutionary pressure acting on the HVR-2 of G protein of Saudi HRSV-B strains, the ratio between non-synonymous and synonymous mutations (dN/dS or ω) was estimated using SNAP program. The dN/dS ratio of Saudi strains ranged from 0.121 (between Riyadh 52/2008 and Riyadh 86/2009) to 1.188 (between Riyadh 87/2009 and Riyadh 133/2009), with a mean value of 0.454. A ratio of one indicates neutral selection, while ratios greater than one implies positive (Darwinian) selection and less than one implies purifying selection. These results indicate that although most of Saudi strains have a stabilizing pressure on G protein, evidences of positive selection on certain Saudi strains exist, which may constitute a signal of rapid evolutionary process in the future.

Discussion

Variability of HRSV is a principle factor that mediates the widespread distribution of the virus and the ability to induce recurrent infections. Two main groups and several sub-group genotypes were identified on the basis of antigenic and molecular diversity of several HRSV genes/proteins; principally G [6, 8]. Viruses belonging to different HRSV groups and genotypes are continuously co-circulating during individual epidemics with a predominance shift each season as a result of alternating and incomplete herd immunity [5]. Information about genetic variability of HRSV strains, particularly group B, in Middle East countries is extremely limited. In this report, the diversity of HRSV-B strains circulating among Saudi population of Riyadh was analyzed and compared to the major genotypes identified worldwide.

All Saudi HRSV-B strains were classified as belonging to BA genotype by sequence and phylogenetic analysis of the C-terminal portion of G protein HVR-2 (Fig. 1). This genotype was first reported in Buenos Aires (BA) at 1999 [16] and was characterized by 60-nucleotide duplication in the HVR-2 of G gene. A later report indicated that the origin of this genotype may return back to 1998 when the virus was detected in an archival sample collected from Spain [17]. During the last decade, the BA genotype was continuously recorded in many countries worldwide; Belgium [24], Japan [25], Kenya [26], India [27], Canada, UK [28], Brazil [29], China [9], Iran [30], Korea [10], and Croatia [31]. The collective epidemiological data indicated that BA viruses are rapidly disseminating globally to the extent that they began gradually to replace other HRSV-B genotypes. None of these reports identified non-BA genotypes after the epidemic season 2005–2006. Although the exact advantage of genotype BA is unknown, it was suggested that two mechanisms may play a role in the evolution and spread of this genotype: (1) a change in the antigenic structure of G protein due to sequence duplication, particularly with respect to O-glycosylation sites [32]. This change may lead to improve virus survival in nature, (2) the immature immune state of individuals towards the newly evolved genotype [28].

Since its emergence, BA genotype experienced sequence divergence due to continuous accumulation of point mutations in the sequence of duplicated region. This variation permitted further subdivision of BA genotype

into subgroups BA-I to BA-X [17, 33] and CB-B [10]. The dynamic replacement of these subgroups in various parts of the world has lead to extinction of BA-I, II, and III subgroups by the years 2002, 2005 and 2005, respectively. Subgroup BA-IV predominated during the years 2004–2008, and then replaced, to a significant extent, by subgroups VII-X [17]. Phylogenetic analysis of Saudi strains showed marked temporal correlation with the prevalent subgroups worldwide; BA-IV for 2008 strains and BA-CB-B and BA-X for 2009 strains (Fig. 1).

It is worth to mention that Saudi strains were related chronologically to those identified in Korea during the same epidemic season 2008–2009 [10]. Both Saudi and Korean strains that belonged to the subgroup BA-X were recovered early during 2009 (January and February, respectively). A novel BA subgroup (CB-B) was identified for the first time in Korea during October and November 2009. CB-B subgroup was not detected again in any district of the world except in this study, where it was identified in samples recovered at January 2009. This result may indicate that such subgroup was originated few months before being detected in Korea; potentially in the Middle East region.

The genetic diversity of HRSV-B strains is mostly attributed to two principle changes observed in the G protein: (1) extensive variation in the protein length (13 different amino acid lengths were recorded) as a result of sequence insertions, deletions and duplication, and the usage of different stop codons [24]; (2) amino acid substitutions that may be genotype-specific [8, 15]. In Saudi strains, a significant homogeneity in the protein length (312 aa) and the usage of stop codon (TAA) was demonstrated. The same observation was recorded for the predominant BA subgroups circulating worldwide since 2008; BA-IV, IX, X (Fig. 2) [10, 31, 33]. Whether these criteria become dominant for HRSV-B circulating globally in the next epidemics, it needs close monitoring and further investigation.

All genetic changes recorded in Saudi strains appeared in the form of base substitutions. Among 18 mutations recorded, only five were specific for individual Saudi strains, while the majority was described before for other members of BA genotype (Fig. 2). It is important to mention that a significant number of base substitutions (no = 9) occurred in the duplicated region. These mutations, in particular, have led to a significant change in the glycosylation pattern evidenced by the appearance of a unique N-glycosylation site in Riyadh 28/2008 strain and the alternative existence/absence of several O-glycosylation sites in Saudi strains (at residues 245, 246, 267, 270, 275). Modulation of the number and distribution of the glycosylation sites can affect the expression of certain epitope(s) either by masking the antigenic sites or influencing the antibody recognition [34]. Therefore, these results may indicate that the duplicated region is still under evolutionary pressure and new BA subgenotypes can be expected in the future.

In conclusion, this is the first report that describes and shed light on the genetic variability of HRSV-B strains in Saudi Arabia. The study extends the notion that BA genotype is expanding to involve new regions of the world with same extent and evolutionary process. Since we did not analyze large number of HRSV-B strains on chronological and/or geographical bases, the presented data provide primary profile on how HRSV-B distributed in Saudi Arabia. Additional isolates from different districts of the kingdom in consecutive epidemic seasons should be included in the future studies for more comprehensive evaluation of virus circulation.

Acknowledgments This study was partially supported by the Research Grant (RGP-VPP-136), Deanship of Scientific Research, King Saud University.

Conflicts of interest The authors declare that they have no conflicts of interest.

References

1. P.A. Cane, Rev. Med. Virol. **11**, 103–116 (2001)
2. A.R. Falsey, E.E. Walsh, Clin. Microbiol. Rev. **13**, 371–384 (2000)
3. P.L. Collins, J.E. Crowe, in *Fields Virology*, ed. by D.M. Knipe, P.M. Howley (Lippincott Williams and Wilkins, Philadelphia, 2007), pp. 1601–1646
4. L.J. Anderson, J.C. Hierholzer, C. Tsou, R.M. Hendry, B.F. Fernie, Y. Stone, K. McIntosh, J. Infect. Dis. **151**, 626–633 (1985)
5. W.M. Sullender, Clin. Microbiol. Rev. **13**, 1–15 (2000)
6. P.R. Johnson, M.K. Spriggs, R.A. Olmsted, P.L. Collins, Proc. Natl. Acad. Sci. USA. **84**, 5625–5629 (1987)
7. P. Rueda, T. Delgado, A. Portela, J.A. Melero, B. Garcia-Barreno, J. Virol. **65**, 3374–3378 (1991)
8. T.C. Peret, C.B. Hall, G.W. Hammond, P.A. Piedra, G.A. Storch, W.M. Sullender, C. Tsou, L.J. Anderson, J. Infect. Dis. **181**, 1891–1896 (2000)
9. Y. Zhang, W. Xu, K. Shen, Z. Xie, L. Sun, Q. Lu, C. Liu, G. Liang, J.A. Beeler, L.J. Anderson, Arch. Virol. **152**, 1425–1434 (2007)
10. Y.H. Baek, E.H. Choi, M.S. Song, P.N. Pascua, H.I. Kwon, S.J. Park, J.H. Lee, S.I. Woo, B.H. Ahn, H.S. Han, Y.S. Hahn, K.S. Shin, H.L. Jang, S.Y. Kim, Y.K. Choi, Arch. Virol. **157**, 1039–1050 (2012)
11. W.B. Coggins, E.J. Lefkowitz, W.M. Sullender, J. Clin. Microbiol. **36**, 3552–3557 (1998)
12. K.M. McConnochie, C.B. Hall, E.E. Walsh, K.J. Roghmann, J. Pediatr. **117**, 52–62 (1990)
13. T.C. Peret, C.B. Hall, K.C. Schnabel, J.A. Golub, L.J. Anderson, J. Gen. Virol. **79**, 2221–2229 (1998)
14. A. Arnott, S. Vong, S. Mardy, S. Chu, M. Naughtin, L. Sovann, C. Buecher, J. Beaute, S. Rith, L. Borand, N. Asgari, R. Frutos, B. Guillard, S. Touch, V. Deubel, P. Buchy, J. Clin. Microbiol. **49**, 3504–3513 (2011)

15. M. Venter, S.A. Madhi, C.T. Tiemessen, B.D. Schoub, J. Gen. Virol. **82**, 2117–2124 (2001)
16. A. Trento, M. Galiano, C. Videla, G. Carballal, B. Garcia-Barreno, J.A. Melero, C. Palomo, J. Gen. Virol. **84**, 3115–3120 (2003)
17. A. Trento, I. Casas, A. Calderon, M.L. Garcia-Garcia, C. Calvo, P. Perez-Brena, J.A. Melero, J. Virol. **84**, 7500–7512 (2010)
18. F.N. Al-Majhdi, A. Al-Jarallah, M. Elaeed, A. Latif, L. Gissmann, H.M. Amer, Int. J. Virol. **5**, 154–163 (2009)
19. F.N. Almajhdi, M.A. Farrag, H.M. Amer, Arch. Virol. (2013). doi:10.1007/s00705-013-1792-6
20. M. Nei, T. Gojobori, Mol. Biol. Evol. **3**, 418–426 (1986)
21. B. Korber, in *Computational Analysis of HIV Molecular Sequences*, ed. by A.G. Rodrigo, G.H. Learn (Kluwer Academic Publishers, Dordrecht, 2000), pp. 55–72
22. K. Tamura, D. Peterson, N. Peterson, G. Stecher, M. Nei, S. Kumar, Mol. Biol. Evol. **28**, 2731–2739 (2011)
23. G.W. Wertz, P.L. Collins, Y. Huang, C. Gruber, S. Levine, L.A. Ball, Proc. Natl. Acad. Sci. USA. **82**, 4075–4079 (1985)
24. K.T. Zlateva, P. Lemey, E. Moes, A.M. Vandamme, M. van Ranst, J. Virol **79**, 9157–9167 (2005)
25. Y. Kuroiwa, K. Nagai, L. Okita, I. Yui, T. Kase, T. Nakayama, H. Tsutsumi, J. Med. Virol. **76**, 241–247 (2005)
26. P.D. Scott, R. Ochola, M. Ngama, E.A. Okiro, D.J. Nokes, G.F. Medley, P.A. Cane, J. Med. Virol. **74**, 344–354 (2004)
27. S. Parveen, S. Broor, S.K. Kapoor, K. Fowler, W.M. Sullender, J. Med. Virol. **78**, 659–665 (2006)
28. A. Trento, M. Viegas, M. Galiano, C. Videla, G. Carballal, A.S. Mistchenko, J.A. Melero, J. Virol. **80**, 975–984 (2006)
29. V.F. Botosso, P.M. Zanotto, M. Ueda, E. Arruda, A.E. Gilio, S.E. Vieira, K.E. Stewien, T.C. Peret, L.F. Jamal, M.I. Pardini, J.R. Pinho, E. Massad, O.A. Sant'anna, E.C. Holmes, E.L. Durigon, PLoS Pathog. **5**, e1000254 (2009)
30. E. Faghihloo, V. Salimi, F. Rezaei, M. Naseri, S. Mamishi, M. Mahmoodi, T. Mokhtari-Azad, Iran. J. Pediatr. **21**, 58–64 (2011)
31. D. Forcic, J. Ivancic-Jelecki, G. Mlinaric-Galinovic, G. Vojnovic, A. Babic-Erceg, I. Tabain, J. Med. Virol. **84**, 1985–1992 (2012)
32. M. Sato, R. Saito, T. Sakai, Y. Sano, M. Nishikawa, A. Sasaki, Y. Shobugawa, F. Gejyo, H. Suzuki, J. Clin. Microbiol **43**, 36–40 (2005)
33. I.C. Dapat, Y. Shobugawa, Y. Sano, R. Saito, A. Sasaki, Y. Suzuki, A. Kumaki, H. Zaraket, C. Dapat, T. Oguma, M. Yamaguchi, H. Suzuki, J. Clin. Microbiol. **48**, 3423–3427 (2010)
34. J.A. Melero, B. Garcia-Barreno, I. Martinez, C.R. Pringle, P.A. Cane, J. Gen. Virol. **78**, 2411–2418 (1997)
35. A. Roca, M.P. Loscertales, L. Quinto, P. Perez-Brena, N. Vaz, P.L. Alonso, J.C. Saiz, J. Gen. Virol. **82**, 103–111 (2001)
36. A. Blanc, A. Delfraro, S. Frabasile, J. Arbiza, Arch. Virol. **150**, 603–609 (2005)
37. R.A. Karron, D.A. Buonagurio, A.F. Georgiu, S.S. Whitehead, J.E. Adamus, M.L. Clements-Mann, D.O. Harris, V.B. Randolph, S.A. Udem, B.R. Murphy, M.S. Sidhu, Proc. Natl. Acad. Sci. USA. **94**, 13961–13966 (1997)
38. W.M. Sullender, M.A. Mufson, L.J. Anderson, G.W. Wertz, J. Virol. **65**, 5425–5434 (1991)